Prenatal characterization of a novel inverted *SMAD2* duplication by mate-pair sequencing in a fetus with dextrocardia

Cinthya Zepeda-Mendoza[1], Anna Essendrup[2], Stephanie Smoley[2], Sarah Johnson[3], Nicole Hoppman[2], George Vasmatzis[3], Daniel Jackson[4], Hutton Kearney[2], and Linda Baughn[2]

[1]ARUP Laboratories
[2]Mayo Clinic Department of Laboratory Medicine and Pathology
[3]Mayo Clinic Center for Individualized Medicine
[4]University of Missouri System

October 1, 2020

Abstract

A fetus harboring a duplication of *SMAD2* exons 1-6 presented with dextrocardia and pulmonary hypoplasia. Mate-pair sequencing revealed the duplication to be in inverted tandem orientation to the wild-type *SMAD2* allele, disrupting its sequence and decreasing expression. These observations suggest *SMAD2* to be responsible for the fetal dextrocardia.

Key clinical message

This case report underlines the importance of molecular characterization of genomic duplications and other structural variants in the prenatal setting to guide clinical interpretation, genetic counseling, and perinatal medical care.

Introduction

Accurate characterization of genomic structural variation is essential to prenatal genetic screening and clinical diagnosis of fetuses with abnormal ultrasound findings. Currently, prenatal genomic abnormalities ranging from aneuploidy to single-nucleotide variants (SNVs) can be detected with a combination of genetic testing strategies, including but not limited to chromosome karyotyping, fluorescence in-situ hybridization (FISH), chromosome microarray (CMA), and next-generation sequencing (NGS) [1-4]. CMA in particular has been shown to substantially increase prenatal diagnostic yields compared to chromosome banding analyses, given its ability to detect clinically relevant submicroscopic duplications and deletions (copy-number variants, CNVs) [1, 5]. Chromosome deletions have been extensively associated with abnormal phenotypes through their impact on haploinsufficient gene expression, an observation that facilitates prenatal deletion classification and reporting. In contrast, fetal chromosome duplications have posed a more challenging and potentially uncertain interpretative context: at the functional level, duplications have the potential to result in overexpression of triplosensitive genes, disruption of haploinsufficient genes at their breakpoints, or creation of gene fusions, all of which could contribute to pathogenesis [6]. The clinical utility of characterizing duplications by NGS to inform their mechanism of pathogenesis has been previously shown [6], with duplications in inverted orientation being more often associated with complex chromosome rearrangements.

Mate pair sequencing (MPseq) is an NGS technology specifically designed for the detection of genomic structural variants [7, 8]. The method is based on the generation of large insert libraries (2-5kb) followed by paired-end (PE) sequencing. Mapping of PE reads can easily reveal the presence of structural rearrangements

1

with a high degree of confidence, including translocations, inversions, insertions, and copy-number variants [8-10].

In this report, we describe the MPseq analysis of a fetus with dextrocardia, pulmonary hypoplasia, and a complex conotruncal anomaly. Dextrocardia is a rare congenital condition in which the heart points towards the right side of the chest rather than the left [11] and is frequently associated with heterotaxy, with important morbidity contributions due to abnormal pulmonary venous connection and ventricle obstructions [12]. CMA performed on fetal amniotic fluid revealed a 145 Kbp duplication at 18q21.1, including exons 1-6 of the SMAD Family Member 2 (*SMAD2*) gene. While *SMAD2* plays an important role in mouse left-right embryonic patterning [13] and mutations have been linked to dextrocardia [14] and other heart abnormalities in humans [15], the CMA results alone could not be used to determine if the duplicated segment disrupted *SMAD2* gene function. MPseq was necessary to clarify the role of this duplication in the clinical phenotype of the fetus by revealing the duplicated segment to be in inverted tandem orientation to the normal chromosome segment. This finding was predicted to disrupt *SMAD2* as evidenced by its reduced gene expression in cultured amniocytes.

Results

Case presentation

A 22-year-old P3G4 woman (M1) was referred for genetic testing due to fetal anomalies detected by ultrasound. The fetus (the proband in this report, D4) was found to have dextrocardia, conotruncal anomaly, and pulmonary hypoplasia resulting in total anomalous pulmonary venous return (TAPVR). Previous M1 pregnancies include a phenotypically normal child (D1), one miscarriage (D2), and a child with heterotaxy born with a single ventricle, who passed away the day after delivery (D3) (Figure 1). During the writing of this report, M1 had a fifth pregnancy complicated with fetal pericardial effusion and premature atrial contractions; the child (D5) was born at 37 weeks gestation and is alive and well at one month. M1 presents with sickle cell trait, but is otherwise reportedly healthy, without major surgeries or hospitalizations between pregnancies, and no evidence of heterotaxy syndrome as revealed by a recent echocardiogram. Maternal and paternal families have no known consanguinity.

To characterize the etiology of the fetal dextrocardia found in proband (D4), genetic testing was performed on an amniocentesis sample from M1. Aneuploidy FISH analysis revealed a normal signal pattern for chromosomes X, Y, 13, 18, and 21 (Supplementary Figure 1), however CMA analyses revealed the presence of a duplication at 18q21.1 at g.47858016_48002991 (human genome version GRch38). The duplication, approximately 145 Kbp in size, partially overlapped *SMAD2*(NM_001003652) from exons one through six (Figure 2A). Subsequent CMA studies on blood of the phenotypically unaffected mother (M1) revealed the duplication to be inherited.

SMAD proteins regulate cell growth and differentiation by mediating TGFβ signaling [16]. In humans, heterozygous loss-of-function mutations of *SMAD2* have been linked to congenital heart disease, including dextrocardia [14] and arterial aneurysms and dissections [15]. While *SMAD2* has a highly significant genotype-phenotype association in proband, the interpretation of the duplication by CMA alone was uncertain given that CMA does not give positional information for the rearrangement (i.e. the duplication could be located in direct or inverted orientation to the normal *SMAD2* allele, or it could be located elsewhere in the genome). To clarify the contribution of the *SMAD2* duplication in proband's phenotype, MPseq analysis was performed on fetal amniocytes. The duplication was mapped to chr18:47,850,720-48,005,918 (hg38). PCR reactions amplified the predicted MPseq junctions and breakpoints are estimated to be located between chr18:47850012-48005625 and chr18:47851237-47852159 (hg38) (Supplementary Figure 2). Structurally, the duplication was found to be in inverted tandem orientation to the normal *SMAD2* allele, a pattern predicted to disrupt its coding sequence (Figure 2B, 2C, and Supplementary Figure 3). Gene expression analysis of the partial *SMAD2* duplication revealed *SMAD2* mRNA levels to be reduced by half compared to a normal male fetal amniocyte control, in agreement with the predicted *SMAD2* allele disruption (Figure 3).

Discussion

Clinical interpretation of genomic duplications identified by CMA is a challenge faced by clinical genetic laboratories. As exemplified by this report, clarifying the genotype-phenotype correlations of genomic duplications requires molecular breakpoint characterization since duplications can be pathogenic by more than one mechanism, including gene dosage alterations, gene disruption at breakpoints, creation of pathogenic gene fusions, and even the re-wiring of the local and distal regulatory landscape of chromatin organization [6, 17].

For the presented prenatal case, the 145 Kbp duplication was found to be in inverted tandem orientation to the normal *SMAD2* allele, disrupting the coding sequence of this gene and reducing *SMAD2* expression to about half its normal value. The characterization of this duplication and its associated functional impact could not have been achieved without a molecular method, such as MPseq, to determine rearrangement positional information. Similar to our case, MPseq has been applied to the analysis of structural variants in the prenatal setting, and shown to resolve complex chromosome rearrangements with important consequences to prenatal diagnosis and perinatal management [18]. To our knowledge, this is the first report of a partial *SMAD2* duplication in inverted orientation associated with heterotaxy. In a murine model, *Smad2* has been shown to act in a dosage-dependent manner in the Nodal signaling pathway to regulate left-right developmental patterning [13, 19]. Similarly in humans, *SMAD2* has been classified as a haploinsufficient gene as evidenced by reports of deletions of *SMAD2* in association with congenital heart defects, including dextrocardia [14].

It remains uncertain how the presence of the duplication in proband's mother was not tied to any major clinical concerns despite the inheritance of the duplication in affected offspring. While duplications are often parentally inherited, the functional outcome and interpretation can be obscured by mosaicism, incomplete penetrance, and/or variable expressivity. For example, a subject with a parentally inherited 1.68 Mb deletion encompassing *SMAD2* reported in the Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources (DECIPHER)[21] does not present with congenital heart defects, but pulmonary stenosis was detected, a feature which can also be seen in heterotaxy (Supplementary Table 2). In contrast, ClinVar [22] entries of *SMAD2* mutations include four cases with reported cardiac anomalies (Supplementary Table 3) [14]. For M1 we were unable to determine if the partial *SMAD2* duplication was mosaic, or if additional heterotaxy-related pathogenic variants were present in her genome or that of her partners.

With more cases being published and variants made accessible to the genetics community through collective efforts such as the DECIPHER and ClinVar databases, prediction of clinical outcomes of prenatal duplications may be more accurate. Nucleotide-level resolution studies will, ultimately, provide the final answer to the complex problem of clinical duplication interpretation in routine genetic testing.

Materials and Methods

Genetic testing was performed on an amniocentesis sample from M1. Aneuploidy FISH analysis was performed on 100 nuclei from uncultured amniocytes and tested chromosomes X, Y, 13, 18, and 21. CMA was performed on cultured amniocytes using the Affymetrix CytoScan HD platform according to manufacturer' instructions. MPseq was performed on DNA extracted from amniocytes using AutoPure LS (Qiagen, Hilden, Germany). MPseq libraries were prepared using Nextera Mate Pair Library Preparation Kit (Illumina, San Diego, CA), which were subsequently purified and used for short-read library preparation using TruSeq DNA Library Prep kit (Illumina, San Diego, CA). Purified libraries were sequenced in an Illumina HiSeq 2500 using RapidRun mode to obtain 101bp reads. MPseq data was processed with BIMAv3[10] and analyzed with SVAtools version 0.24.9 [8]. Polymerase chain reaction (PCR) experiments were performed to amplify breakpoint junctions identified by MPseq. Gene expression experiments were performed on total mRNA purified from cultured amniocytes of the proband and a normal control male with normal *SMAD2* copy number and no phenotypic evidence of dextrocardia. Relative mRNA levels were determined by real-time RT-PCR analysis using Applied Biosystems TaqMan Assays (Thermo Fisher Scientific, Waltham, MA) targeting the *SMAD2* exon junctions 2/3, 4/5, 6/7, and 9/10. Three technical replicates were used to calculate standard deviation for both the proband and control RNA expression (primers used are listed in Supplementary Table 1).

3

Ethical Approval

Funding

Author Contributions

Conceptualization, CJZM, HMK, and LBB; Formal analysis, CJZM, AE, SAS, SHJ, NLH, HMK, LBB; Writing, CJZM and LBB; Resources, DLJ, GV, LBB

Conflict of Interest

The authors declare that there is no conflict of interest

References

1. *Committee Opinion No. 581: the use of chromosomal microarray analysis in prenatal diagnosis.* Obstet Gynecol, 2013. **122** (6): p. 1374-7.

2. *Practice Bulletin No. 162: Prenatal Diagnostic Testing for Genetic Disorders.* Obstet Gynecol, 2016. **127** (5): p. e108-22.

3. *Practice Bulletin No. 163: Screening for Fetal Aneuploidy.*Obstet Gynecol, 2016. **127** (5): p. e123-37.

4. Manegold-Brauer, G., S. Hahn, and O. Lapaire, *What does next-generation sequencing mean for prenatal diagnosis?* Biomark Med, 2014. **8** (4): p. 499-508.

5. Lovrecic, L., et al., *Clinical utility of array comparative genomic hybridisation in prenatal setting.* BMC Med Genet, 2016.**17** (1): p. 81.

6. Newman, S., et al., *Next-generation sequencing of duplication CNVs reveals that most are tandem and some create fusion genes at breakpoints.* Am J Hum Genet, 2015. **96** (2): p. 208-20.

7. Talkowski, M.E., et al., *Next-generation sequencing strategies enable routine detection of balanced chromosome rearrangements for clinical diagnostics and genetic research.* Am J Hum Genet, 2011.**88** (4): p. 469-81.

8. Johnson, S.H., et al., *SVAtools for junction detection of genome-wide chromosomal rearrangements by mate-pair sequencing (MPseq).*Cancer Genet, 2018. **221** : p. 1-18.

9. Smadbeck, J.B., et al., *Copy number variant analysis using genome-wide mate-pair sequencing.* Genes Chromosomes Cancer, 2018.**57** (9): p. 459-470.

10. Drucker, T.M., et al., *BIMA V3: an aligner customized for mate pair library sequencing.* Bioinformatics, 2014. **30** (11): p. 1627-9.

11. Oztunc, F., et al., *Diagnosis and outcome of pregnancies with prenatally diagnosed fetal dextrocardia.* J Matern Fetal Neonatal Med, 2015. **28** (9): p. 1104-7.

12. Wang, X., et al., *Comparing levocardia and dextrocardia in fetuses with heterotaxy syndrome: prenatal features, clinical significance and outcomes.* BMC Pregnancy Childbirth, 2017.**17** (1): p. 393.

13. Nomura, M. and E. Li, *Smad2 role in mesoderm formation, left-right patterning and craniofacial development.* Nature, 1998.**393** (6687): p. 786-90.

14. Zaidi, S., et al., *De novo mutations in histone-modifying genes in congenital heart disease.* Nature, 2013. **498** (7453): p. 220-3.

15. Micha, D., et al., *SMAD2 Mutations Are Associated with Arterial Aneurysms and Dissections.* Hum Mutat, 2015. **36** (12): p. 1145-9.

16. Stroschein, S.L., et al., *Negative feedback regulation of TGF-beta signaling by the SnoN oncoprotein.* Science, 1999.**286** (5440): p. 771-4.

17. Lupianez, D.G., et al., *Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions.* Cell, 2015. **161** (5): p. 1012-1025.

18. Ordulu, Z., et al., *Structural Chromosomal Rearrangements Require Nucleotide-Level Resolution: Lessons from Next-Generation Sequencing in Prenatal Diagnosis.* Am J Hum Genet, 2016. **99** (5): p. 1015-1033.

19. Schier, A.F., *Nodal signaling in vertebrate development.* Annu Rev Cell Dev Biol, 2003. **19** : p. 589-621.

20. Maldjian, P.D. and M. Saric, *Approach to dextrocardia in adults: review.* AJR Am J Roentgenol, 2007. **188** (6 Suppl): p. S39-49; quiz S35-8.

21. Firth, H.V., et al., *DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources.* Am J Hum Genet, 2009. **84** (4): p. 524-33.

22. Landrum, M.J., et al., *ClinVar: public archive of interpretations of clinically relevant variants.* Nucleic Acids Res, 2016. **44** (D1): p. D862-8.

Figure legends:

Figure 1. Pedigree showing segregation of a 145 Kbp duplication at 18q21.1, including exons 1-6 of *SMAD2* . Proband is indicated with a black arrow (D4).

Figure 2. arr[GRCh37] 18q21.1(45384387_45529362)x3 mat duplication in proband. A) Screenshot of the duplication position as seen in ChAS software. The duplication is shown in the first horizontal panel as a blue rectangle. Probe positions are shown in the middle panel as dots. Notice the elevated signal for probes directly below the blue rectangle, which were used for calling the duplication positions. Gene positions (*SMAD2* and *ZBTB7C*) are shown in the third panel, along with the nucleotide positions in chromosome 18. Exon numbers are displayed for *SMAD2* . The shaded rectangle encompasses*SMAD2* positions, and shows the partial duplication. The 145 Kbp duplication overlapped exons one through six of *SMAD2* , which could potentially disrupt the gene's function. In addition, the duplication was found to be maternally inherited in proband, but the mother was phenotypically normal. B) The *SMAD2* rearrangement as seen in our MPseq analysis pipeline. This junction plot displays four breakpoints representing the inversion event (colored dots indicate orientation of mapped reads, with blue corresponding to forward strand and red the reverse strand). Notice the presence of the blue-blue and red-red dot associations predicted to be seen in a duplication with inverted orientation. Only discordant mate-pair reads with end positions mapping >15 kb are shown. C) Schematic representation of the *SMAD2* duplication seen in proband compared to the reference allele. Concordant and discordant mate-pair fragments and their corresponding orientations are shown in colored text.

Figure 3. *SMAD2* RNA expression analysis. Total mRNA was isolated from cultured proband amniocytes, and RNA extracted from the amniocytes of a normal male with normal *SMAD2* copy number and no phenotypic evidence of dextrocardia was used as a control. Real-time RT-PCR analysis targeted the *SMAD2* exon junctions 2/3, 4/5, 6/7, and 9/10 indicated underneath the black horizontal bars in both the proband (grey bars and Pt label) and the control (black bars and Ctrl label). NTC: no template control. The expression of *SMAD2* is significantly reduced in proband compared to the normal control.

I

□—#—○ M1 □   □
        24 y

II

◇   △   ⊠   ◉   ○
D1   D2   D3   D4   D5
? y   16 weeks   1 day   2 y   1 mo

◕ Heterotaxy   ◉ *SMAD2*
                        duplication